NORTH-HOLLAND

A Note on the Davis-Kahan $\sin\Theta$ Theorem

Jianjun Xie
Institute of Mathematics
Fudan University
Shanghai 200433, P. R. China

Submitted by Hans Schneider

ABSTRACT

By using a series of inequalities for singular values of matrix products, we obtain perturbation theorems for invariant subspaces of Hermitian matrices, which are sharper than the second Davis-Kahan $\sin\Theta$ theorem.

1. INTRODUCTION

Assessing the accuracy of an approximate invariant subspace in terms of a residual is an important problem in matrix computations. For Hermitian matrices the existence of a perturbed invariant subspace is often obvious. Thus we may assume the existence of the perturbed invariant subspace and proceed directly to bounds on the canonical angles between the original subspaces and its perturbation. This general approach is due to Davis and Kahan [2]. The following theorem due to Davis and Kahan is called the second $\sin\Theta$ theorem.

THEOREM 1. *Let $A \in \mathbb{C}^{n\times n}$ and $M \in \mathbb{C}^{k\times k}$ be Hermitian, and A have the spectral resolution*

$$\begin{pmatrix} X_1^H \\ X_2^H \end{pmatrix} A(X_1, X_2) = \operatorname{diag}(L_1, L_2),$$

655 Avenue of the Americas, New York, NY 10010

0024-3795/97/$17.00
PII S0024-3795(96)00207-8

where (X_1, X_2) *is unitary with* $X_1 \in \mathbb{C}^{n\times k}$. *Here* ^{H} *takes the conjugate transpose. Let* $Z \in \mathbb{C}^{n\times k}$ *have orthonormal columns, and let* $R = AZ - ZM$. *For some* $\delta > 0$ *suppose that*

$$\lambda(M) \subseteq [\alpha, \beta],$$

$$\lambda(L_2) \subseteq (-\infty, \alpha - \delta] \cup [\beta + \delta, +\infty)$$

(*or vice versa*). *Then for any unitarily invariant norm*

$$|||\sin \Theta[R(X_1), R(Z)]||| \leqslant \frac{|||R|||}{\delta},$$

where $\lambda(M)$ *denotes the set of all eigenvalues of* M, $R(X_1)$ *denotes the column space of* X_1, *and* $|||\cdot|||$ *denotes any unitarily invariant norm.*

It is well known that $|||X_2^H Z||| = |||\sin \Theta[R(X_1), R(Z)]|||$ (see [2, 5]). As the singular values of $X_2^H Z$ are the sines of the individual canonical angles between $R(X_1)$ and $R(Z)$, it is important to give tight upper bounds for sums of some of these singular values.

If we know more information about the spectra of L_2 and M, we can obtain sharper results than those in Theorem 1.

2. DEFINITIONS AND LEMMAS

First, we give two definitions:

DEFINITION 1. Let $M \in \mathbb{C}^{k\times k}$ and $L_2 \in \mathbb{C}^{(n-k)\times(n-k)}$ be Hermitian. Let the eigenvalues of M and L_2 be $\lambda_1, \lambda_2, \dots, \lambda_k$ and $\mu_1, \mu_2, \dots, \mu_{n-k}$. For some $\delta > 0$, suppose that $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_k$, $\mu_i \in (-\infty, \lambda_1 - \delta] \cup [\lambda_k + \delta, +\infty)$, $i = 1, 2, \dots, n-k$ (or vice versa) and

$$\lambda_i' = \lambda_i - \frac{\lambda_k + \lambda_1}{2}, \qquad i = 1, 2, \dots, k,$$

$$\mu_i' = \mu_i - \frac{\lambda_k + \lambda_1}{2}, \qquad i = 1, 2, \dots, n-k,$$

so $|\lambda_i'| \leqslant (\lambda_k - \lambda_1)/2$, $|\mu_i'| > (\lambda_k - \lambda_1)/2$. We may assume without loss of generality that

$$|\lambda_1'| \geqslant |\lambda_2'| \geqslant \cdots \geqslant |\lambda_k'| \quad \text{and} \quad |\mu_1'| \leqslant |\mu_2'| \leqslant \cdots \leqslant |\mu_{n-k}'|.$$

Then d_i are *absolute separations* of the spectra of M and L_2, if

$$d_i = |\mu_i'| - |\lambda_i'|, \qquad i = 1, 2, \dots, l,$$

where $l = \min(k, n-k)$.

DEFINITION 2. Let $M \in \mathbb{C}^{k\times k}$ and $L_2 \in \mathbb{C}^{(n-k)\times(n-k)}$ be Hermitian. Let the eigenvalues of M and L_2 be

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k \quad \text{and} \quad \mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_{n-k}.$$

For some $\delta > 0$, suppose that $\mu_i \in [\, \lambda_1 + \delta, +\infty)$ (or vice versa). Then D_i are *general separations* of the spectra of M and L_2 if

$$D_i = |\mu_i - \lambda_i|, \qquad i = 1, 2, \dots, l.$$

From the two definitions we know

$$\delta \leqslant d_1 \leqslant d_2 \leqslant \cdots \leqslant d_l,$$

$$\delta \leqslant D_1 \leqslant D_2 \leqslant \cdots \leqslant D_l.$$

To establish the main results in this paper, we need the following lemmas.

LEMMA 1 (Fan). *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ have singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n$ and $\tau_1 \geqslant \tau_2 \geqslant \cdots \geqslant \tau_n$. In order for*

$$|||A||| \leqslant |||B|||$$

for every unitarily invariant norm it is necessary and sufficient that

$$\|A\|_{(k)} \leqslant \|B\|_{(k)}, \qquad k = 1, 2, \dots, n,$$

where $\|A\|_{(k)} = \sigma_1 + \sigma_2 + \cdots + \sigma_k$.

The proof can be found in [3, 5].

LEMMA 2 (Horn). *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ have singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n$ and $\tau_1 \geqslant \tau_2 \geqslant \cdots \geqslant \tau_n$. Then*

$$\|AB\|_{(k)} \leqslant \sum_{i=1}^{k} \sigma_i \tau_i, \qquad k = 1, 2, \ldots, n. \tag{1}$$

This result is due to Horn [4].

By using the Birkhoff theorem in [1], Yu proved the following lemma in [6].

LEMMA 3 (Yu). *Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$ have singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n$ and $\tau_1 \geqslant \tau_2 \geqslant \cdots \geqslant \tau_n$. Then*

$$\sum_{i=1}^{k} \sigma_{n-i+1} \tau_i \leqslant \|AB\|_{(k)}, \qquad k = 1, 2, \ldots, n. \tag{2}$$

Lemma 2 and Lemma 3 can be easily extended to the case that A and B are not square matrices. So we have

LEMMA 4. *Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times l}$* (without loss of generality we assume *$m \geqslant n \geqslant l$) have singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n$ and $\tau_1 \geqslant \tau_2 \geqslant \cdots \geqslant \tau_l$. Then*

$$\sum_{i=1}^{k} \sigma_{n-i+1} \tau_i \leqslant \|AB\|_{(k)} \leqslant \sum_{i=1}^{k} \sigma_i \tau_i, \qquad k = 1, 2, \ldots, l. \tag{3}$$

3. PERTURBATION THEOREMS

In this paper, we will establish the following two theorems:

THEOREM 2. *Let $n \times n$ Hermitian matrix A have the spectral resolution*

$$\begin{pmatrix} X_1^H \\ X_2^H \end{pmatrix} A(X_1, X_2) = \operatorname{diag}(L_1, L_2)$$

where (X_1, X_2) is unitary with $X_1 \in \mathbb{C}^{n\times k}$. Let $Z \in \mathbb{C}^{n\times k}$ have orthonormal columns, and for any Hermitian matrix M of order k, let $R = AZ - ZM$. For some $\delta > 0$ suppose that

$$\lambda(M) \subseteq [\alpha, \beta],$$

$$\lambda(L_2) \subseteq (-\infty, \alpha - \delta] \cup [\beta + \delta, +\infty)$$

(or vice versa). Let d_i be absolute separations of the spectra of M and L_2, and let the singular values of $X_2^H Z$ be $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_l$, where $l = \min(k, n-k)$. Then

$$\sum_{i=1}^{m} d_i \sigma_i \leqslant \|R\|_{(m)}, \qquad m = 1, 2, \ldots, l. \tag{4}$$

Proof. Let λ_i, μ_i, λ_i', and μ_i' be as in Definition 1. The substitutions

$$A \leftarrow A - \frac{\lambda_k + \lambda_1}{2} I, \qquad M \leftarrow M - \frac{\lambda_k + \lambda_1}{2} I$$

leave the theorem unchanged, so we may assume without loss of generality that the eigenvalues of M and L_2 are $\lambda_1', \lambda_2', \ldots, \lambda_k'$ and $\mu_1', \mu_2', \ldots, \mu_{n-k}'$ with

$$|\lambda_1'| \geqslant |\lambda_2'| \geqslant \cdots \geqslant |\lambda_k'| \quad \text{and} \quad |\mu_1'| \leqslant |\mu_2'| \leqslant \cdots \leqslant |\mu_{n-k}'|,$$

so $d_i = |\mu_i'| - |\lambda_i'|$, $i = 1, 2, \ldots, l$. Then

$$\begin{aligned} \|R\|_{(m)} &\geqslant \|X_2^H R\|_{(m)} \\ &= \left\| L_2(X_2^H Z) - (X_2^H Z) M \right\|_{(m)} \\ &\geqslant \left\| L_2(X_2^H Z) \right\|_{(m)} - \left\| (X_2^H Z) M \right\|_{(m)}. \end{aligned}$$

From (3) we have

$$\left\| L_2(X_2^H Z) \right\|_{(m)} \geqslant \sum_{i=1}^{m} |\mu_i'| \sigma_i,$$

$$\left\| (X_2^H Z) M \right\|_{(m)} \leqslant \sum_{i=1}^{m} |\lambda_i'| \sigma_i.$$

Hence, (4) holds. This completes the proof. ■

THEOREM 3. *In the notation of Theorem 2, suppose that*

$$\lambda(M) \subseteq [\alpha, \beta],$$

$$\lambda(L_2) \subseteq [\beta + \delta, +\infty)$$

(or vice versa). Let D_i be general separations of the spectra of M and L_2. Then

$$\sum_{i=1}^{m} D_i \sigma_i \leqslant \|R\|_{(m)}, \qquad m = 1, 2, \ldots, l. \tag{5}$$

Proof. Let λ_i, μ_i be as in Definition 2. By translating the spectra of A and M, we may assume without loss of generality that $\lambda_k > 0$. Now the proof is similar to that of Theorem 2. ∎

From Theorem 2 and $d_i \geqslant \delta$ we have

$$\delta \|X_2^H Z\|_{(m)} \leqslant \|R\|_{(m)}, \qquad m = 1, 2, \ldots, l. \tag{6}$$

Now the result in Theorem 1 follows on applying Lemma 1 to (6). Obviously, Theorem 2 is sharper than Theorem 1.

Sometimes, if d_1 is very small, $|||R|||/\delta$ may be large, and from Theorem 1 we can't bound the $|||X_2^H Z|||$ well. But if d_2 is not small and $|||R|||/d_2$ is small, from Theorem 2 we know $|||\operatorname{diag}(0, \sigma_2, \ldots, \sigma_l)||| \leqslant |||R|||/d_2$ is small.

Finally, in the same way as in [2], we can easily extend Theorem 2 and Theorem 3 to the case that the columns of the matrix Z may not be orthonormal.

The author thanks his advisor, Professor Jiang Erxiong, for his careful reading of this paper and helpful comments.

REFERENCES

1 G. D. Birkhoff, Tres observaciones sobre el algebra lineal, *Univ. Nacional Tucuman Rev. Ser. A* 5:147–151 (1946).

2 C. Davis and W. M. Kahan, The rotation of eigenvectors by a perturbation, III, *SIAM J. Numer. Anal.* 7:1–46 (1970).

3 K. Fan, Maximum properties and inequalities for the eigenvalues of completely continuous operators, *Proc. Nat. Acad. Sci. U.S.A.* 37:760–766 (1951).
4 A. Horn, On the singular values of a product of completely continuous operators, *Proc. Nat. Acad. Sci. U.S.A.* 36:374–375 (1950).
5 G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic, Boston, 1990.
6 Y. Yu, Some estimations for the upper bound and lower bound of singular values of matrix products (in Chinese), *Chinese Ann. Math. Ser. A* 6:681–687 (1992).

Received 5 October 1995; final manuscript accepted 12 February 1996